Ades, A. E., Price, M. J., Kounali, D-Z., Akande, V., Wills, G. S., McClure, M. O., Muir, P., & Horner, P. J. (2017). Proportion of Tubal Factor Infertility due to *Chlamydia*: Finite Mixture Modeling of Serum Antibody Titers . *American Journal of Epidemiology*, *185*(2), 124-134. https://doi.org/10.1093/aje/kww117

Proportion of Tubal Factor Infertility due to Chlamydia: finite mixture modeling of serum antibody titers.

AE* Ades, MJ. Price, D. Kounali, VA. Akande, GS. Wills, MO. McClure, P. Muir,

PJ. Horner

(*) Correspondence to: Professor AE Ades; School of Social and Community Medicine, 39 Whatley Road, Bristol BS8 2PS, United Kingdom.

Abstract

This study **examined** whether the proportion of Tubal Factor Infertility (TFI) that is attributable to *Chlamydia trachomatis*, the population excess fraction (PEF), can be estimated from serological data using finite mixture modeling. Whole cell inclusion immune-fluorescence serum antibody titers were recorded in infertile women **who were seen at St. Michael's Hospital, Bristol, between 1985-1995** and classified as TFI cases or controls based on laparoscopic examination. Finite mixture models were used to identify the number of component titer distributions and the proportion of samples in each, from which estimates of PEF were derived. Four titer distributions were identified. The component at the highest titer was found only in samples from women with TFI, but there was also an excess of the second highest titer component in TFI cases. Minimum and maximum estimates of the PEF were 28.0% (95% credible interval: 6.9, 50.0) and 46.8% (95% Credible interval: 23.2, 64.1). Equivalent estimates based on the standard PEF formula from case-control studies were 0% and over 65%. Finite mixture modeling can be applied to serological data to obtain estimates of the proportion of reproductive damage attributable to *Chlamydia trachomatis*. Further studies should be undertaken using modern assays in contemporary, representative populations.

<198 words>

Keywords: Tubal Factor Infertility, *Chlamydia trachomatis*, Population Excess Fraction, finite mixture models, antibody titres.

List of abbreviations

CT	Chlamydia Trachomatis
CT+	CT infected
CT-	Not CT infected
PID	Pelvic Inflammatory Disease
WIF	Whole cell inclusion immune-fluorescence
PEF	Population excess fraction
TFI	Tubal factor fnfertility
OR	Odds ratio
CrI	Credible interval

Chlamydia Trachomatis (CT) is a common sexually transmitted infection of young people which if left untreated will cause pelvic inflammatory disease (PID) in around 16% of women (1). PID may then result in adverse reproductive outcomes such as ectopic pregnancy (EP) or tubal factor infertility (TFI) (2). In spite of research using a wide range of study designs, the precise quantitative relationship between *Chlamydia trachomatis* (CT) and reproductive damage remains elusive (2-4). CT, with or without the development of disease, usually resolves spontaneously. Diagnosed infection is treated, so prospective study of untreated infection is not feasible. The major studies of reproductive outcomes in women with PID (5-7) have been restricted to the small proportion of PID (8) that is diagnosed in hospital. In the 1980s and 1990s large numbers of serological case-control studies were carried out, comparing serum antibody levels in women with PID, EP or infertility, with controls (9-19). These studies invariably showed strong associations between detection of CT antibodies and reproductive damage, but it was difficult to draw quantitative conclusions from them, partly because of confounding between CT and other pathogens also implicated in reproductive morbidity (20), and partly because of the poor, and imprecisely known, sensitivity and specificity of the assays used (21, 22).

In this paper we adopt a new analytic approach to this problem: finite mixture modeling (23). Finite mixture models are used when a distribution, in this case of serum antibody titers, is considered to be a mixture of several components, for example “positives” and “negatives”, and where there is an interest in estimating the proportion of samples in diseased and healthy populations in each component. Finite mixture models are often applied to diagnostic tests which lack a “gold standard”, including to serological data (24-28).

In this paper we apply finite mixture models to a previously published dataset (29). Whole cell inclusion immune-fluorescence (WIF) serum antibody titers were recorded in infertile women classified as having Tubal Factor Infertility (cases) or not (controls) following laparoscopy (Table 1). Note that among the cases, a high proportion of the titers that would normally be considered positive (1:32 and above) are at particularly high levels. This has been observed repeatedly in similar studies of TFI (11, 15, 16, 30, 31), and PID (10, 18). In other words, women at higher risk of reproductive damage are more likely to be CT antibody positive, and are more likely to have particularly high titers than antibody-positive controls.

The Lund studies (5-7) established that clinically diagnosed PID was only associated with reproductive damage in women with laparoscopically confirmed salpingitis. Our analysis, therefore, is premised on the assumption that the exceptionally high positive titers seen in TFI cases reflect an inflammatory reaction to CT that is associated with CT-related salpingitis, and that women with titers at these high levels are at risk of CT-related TFI (32). The purpose of the finite mixture analysis is to determine what proportion of the TFI cases have titers at these high levels.

The estimates of the population excess fraction (PEF) formed in this way will be of substantive interest in the many countries where chlamydia and prevention control strategies are in operation, or being considered. However, in view of the limitations inherent in using data collected many years ago for another purpose, this paper should be seen in part as an exploratory, hypothesis forming, exercise into how and whether finite mixture modeling of anti-CT titers can contribute to an understanding of the role of CT in reproductive damage.

METHODS

Data

The primary dataset consisted of WIF titers from 434 TFI cases confirmed on laparoscopy, and 573 controls who were infertile for other reasons, seen in a Reproductive Medicine Clinic, at St. Michael's hospital in Bristol, between 1985 and 1995. The data were collected in a study exploring the relationship between serum chlamydia antibody titers and detection of tubal damage in infertile women as previously reported (29) (Table 1). Titers of 1:32 or greater would normally be considered positive for CT antibody. Cases had a mean age of 29.3 years (range 18-46) and controls 30.6 years (range 19-47).

A proportion of low titer positives on WIF are likely to be cross-reactions to *Chlamydia pneumonia* (CP) in women with no exposure to CT (33) . A secondary dataset provided additional information on the proportion of CT negatives at each WIF titer. Anonymized samples from women undergoing investigation for infertility during 2013 were submitted to Bristol Public Health Laboratories and tested by WIF at Bristol Public Health Laboratories and by the highly specific Pgp3 CT antibody assay (33) at Imperial College. The analyses reported here concern samples from 301 women with WIF titers at or below 1:1024. Causes of infertility and reproductive outcomes were not recorded.

Models

Three models were examined, each characterized by the number of latent distributions assumed to be present (Table 2). For example, the "2-3" model assumed that the control samples were a mixture of *two* distributions, which we label *CT-* (never infected) and *CT+*

(previously infected no inflammatory response), while TFI samples may come from either of these distributions or from a third distribution, *CT++*, who have had an inflammatory response to CT infection.

Further models were developed when it was found that the "2-3" model did not fit the data. In the "3-3" model a proportion of control samples is also allowed to belong to the *CT++* distribution. In the "3-4" model, a further distribution is proposed, *CT+++*, but only TFI samples may belong to it. These labels should be thought of simply as mnemonics, although as the labels suggest, they are listed in order of increasing titer with *CT-* lowest and representing true CT antibody negatives, and *CT+++* the highest representing exceptionally high levels of serum antibody.

Statistical methods

Finite Mixture Modeling assumes that each distribution G of titers *y* is a mixture of say, *D* underlying latent distributions $f_d(y)$, which we assume are Normal on the log titer scale. We further assume that the only difference between cases (*k*=1) and controls (*k*=0) is in the proportions of samples from each of the latent component distributions, *d=1,…,D*.

$$G_k(y) = \pi_{k1} f_1(y) + \pi_{k2} f_2(y) + \cdots + \pi_{kD} f_D(y)$$

The proportions π_{kd} are the proportion of samples that can be attributed to latent distribution *d*, conditional on case / control status *k*. The means and standard deviations of the component distributions remain the same in cases and controls.

Although titers are reported in categories, they are in fact censored observations on a continuous variable. If we designate the lower boundary of categories {<1:64, 1:64, 1:128, 1:256, 1:512, 1:1024, 1:2028, 1:4096, >1:4096} as: i_{L0} ={- ∞ ,1,2,3,4,5,6,7,8}, and the upper boundaries as i_{Hl} ={1,2,3,4,5,6,7,8,9, ∞ }, we can express the proportion of distribution *d* that falls into the *i*th category as:

$$\theta_{di} = \int_{i_{LO}}^{i_{HI}} f_d(y)dy$$

with *y* on the natural log titer scale. Finally, the proportions of samples, α_{ki} , in each titer category *i* in group *k* is obtained as the inner product of the θ_{di} and the π_{kd} :

$$\alpha_{ki} = \sum_d \pi_{kd}\theta_{di}$$

The observed data is the numbers of samples r_{ki} in each titer category *i* in TFI cases (*k* =1) and controls (*k*=0); this is multi-nomially distributed, with denominators n_k :

$$r_{k,i=1...9} \sim Multinomial(\alpha_{k,1...9}, n_k)$$

This multinomial distribution is the appropriate choice for the observables i.e. *n* trials (samples) with *k* possible outcomes (titers) on each trial, and a fixed probability of each outcome over all the trials.

The secondary data source provides additional information on the proportion of true antibody negatives at each WIF titer (Table 1). This was based on the pgp-3 assay, which we assume to be effectively 100% specific (33). The data in Table 1 provides direct information on the

Field Code Changed

probability ω_{ki} of a sample being in the *CT-* group, conditional on its titer and case / control status. This, in turn, constitute indirect information on θ_{1i} , the probability of a sample having a specified titer, given that it is *CT-*. We can use Bayes Rule to relate the quantities, with the index "1" indicating the *CT-* distribution:

$$\omega_{ki} = \frac{\theta_{1i}\pi_{k1}}{\sum_d \theta_{di}\pi_{kd}}$$

Because we do not know the proportion of TFI cases in the secondary dataset, we further define a weighted average of ω_{1i} for cases and ω_{0i} for controls, with the proportion which are cases p^{TFI} to be estimated from the data:

$$\omega_{.i} = p^{TFI}\frac{\theta_{1i}\pi_{11}}{\sum_d \theta_{di}\pi_{1d}} + (1-p^{TFI})\frac{\theta_{1i}\pi_{01}}{\sum_d \theta_{di}\pi_{0d}}$$

The secondary data r_i providing information on the $\omega_{.i}$ represent the numbers of pgp-3 negatives among the sample n_i with WIF titre *i* (Table 1). These have a Binomial likelihood:

$$r_i \sim Bin(\omega_{.i}, n_i), \quad i = 1,2....6$$

The secondary data covers only the first 6 titer categories, as there are no further pgp-3 negatives in higher WIF categories (Table 1).

Estimates of population excess fraction. In the "2-3" model the distribution with the highest mean titer is only found in the TFI cases. The proportion of TFI samples in the *CT++* distribution in the "2-3" model is therefore a direct estimate of the PEF

Field Code Changed

Field Code Changed

$$PEF^{2-3} = \pi_{1,CT++}$$

In the "3-3" model, a more complicated situation arises: here an estimate of the PEF can be based on the *excess* proportion of samples in the highest *CT++* category in TFI cases *compared to controls*. Thus, it is necessary to take into account that a proportion of the *CT++* observed in the TFI cases would occur anyway, even if TFI status was unrelated to antibody level. Consider the number of *CT++* samples observed in controls, as a proportion of *all CT+* and *CT++*in the controls. If there was no excess *CT++* in the TFI cases we would expect $\pi_{1,CT++}$ to equal:

$$(\pi_{1,CT+} + \pi_{1,CT++})\left(\frac{\pi_{0,CT++}}{\pi_{0,CT+} + \pi_{0,CT++}}\right)$$

The estimate of the PEF is the excess *CT++* in the TFI group, which is therefore the difference between the observed *CT++* in cases, and what we would predict from the controls:

$$PEF^{3-3} = \pi_{1,CT++} - (\pi_{1,CT+} + \pi_{1,CT++})\left(\frac{\pi_{0,CT++}}{\pi_{0,CT+} + \pi_{0,CT++}}\right)$$

For the "3-4" model, we can first follow the same logic as the "2-3" model. The proportion of cases in the *CT+++* group is a direct estimate of the PEF: $PEF^{3-4(1)} = \pi_{1,CT+++}$. This must be considered as a lower bound estimate because it ignores the excess proportion of *CT++* observed in TFI cases.

Acknowledging an excess in *CT++* samples in TFI cases, in addition to the *CT+++*, we can follow the same argument set out for the "3-3" model. This leads to a second estimate of the PEF for "3-4" model, in which we add the proportion in *CT+++* to the excess fraction of the *CT++*:

$$PEF^{3\text{–}4(2)} = \pi_{1,CT+++} + \pi_{1,CT++} - (\pi_{1,CT+} + \pi_{1,CT++})\left(\frac{\pi_{0,CT++}}{\pi_{0,CT+} + \pi_{0,CT++}}\right)$$

This can be considered as an upper bound estimate as it ascribes *all* the excess of *CT++* in cases to a causal effect of CT infection. The equations for each estimate are presented in Table 2. The extent to which these different estimates are vulnerable to confounding is taken up in the discussion.

Results

Model selection

Table 3 compares the predicted titer distributions from each model with the observed data, and shows the goodness of fit (Residual Deviance) at each point. Systematic error is evident in the "2-3" and "3-3" models, in that they *under*-estimate the peak that can be seen in both control and TFI samples at 1:1024, while *over*-estimating the number of samples at 1:512. The "3-4" models on the other hand fit the distribution well at every point.

The fit of models to data is elaborated further in Figure 1 – Figure 4. These plots, depict the fitted component distributions (in color), for each model and separately for cases and controls and are drawn to the correct scale in order to reflect the fitted proportions of each component.

The predicted overall titer distribution (solid black line) is the sum of the components, and can be compared to the observed data shown as a histogram.

In a good fitting model the residual deviance should be no more than the number of data points, which is 8 each in the cases and controls. Therefore the global residual deviance statistics at the foot of Table 3 rule out "2-3" and "3-3" models decisively, with the 3-4 models as an excellent fit.

Parameter estimates and PEF

The estimated mean and standard deviation of the each of the log titer distributions, *CT-*, *CT+*, *CT++*, *CT+++*, are set out in Table 4, along with the proportions of cases and controls in each group. The introduction of a *CT++* distribution for the controls in "3-3" and "3-4" models has the effect of lowering the mean of the *CT+* distribution by about 1 log unit, and somewhat lowering its variance. Similarly, the introduction of a *CT+++* category lowers the mean of the *CT++* group and also reduces its variance.

The secondary data source provides more information on the *CT-* distribution: the mean is raised to a somewhat higher titer, and the variance is reduced. Its main effect is to reduce uncertainty in the means and SDs of the *CT-* and *CT+* distributions. The model fits the secondary data well, with a residual deviance of 4.0 on 6 observations, 2 of which were zeros.

The central estimates of the PEF (Table 4) lie within the range 28%-48%. Although estimates from the "2-3" model can be discounted due its poor fit, it is interesting that it estimates almost exactly the same PEF as the upper bound estimate from the "3-4" model without secondary

data. This shows that when the excess cases in the CT++ group are assumed to cause TFI one obtains very similar results whether or not one distinguishes between *CT++* and *CT+++* distributions.

The secondary data has a slight impact on the estimates of PEF, lowering them by about 4 percentage points. As expected, the probability that a sample in the secondary dataset was from a TFI and not a control was poorly estimated, 0.54 (95% credible interval 0.04, 0.97), barely different from the prior. We consider the estimates from the "3-4" model with secondary data as the best available from the study, due to their greater precision. The secondary data improves identification of the boundary between the *CT-* and *CT+* distributions by eliminating the false positives in *CT+* . This can be observed by comparing the posterior means and 95% credible intervals of the means and standard deviations of the *CT-* and *CT+* distributions (Table 4).

Sensitivity analyses reported in the Web Appendix 3 (Web Table 1, Web Table 2) showed that the main results were robust to reasonable changes in the priors, to distributional assumptions, and to the proportion of TFI cases in the secondary data.

Discussion

Estimates of the proportion of pelvic inflammatory disease, ectopic pregnancy, and infertility that can be attributed to Chlamydia are critical to motivating prevention and control programs for *Chlamydia trachomatis* (CT). A number of authors have attempted to derive estimates from

serological case-control studies (13, 37), but these are confounded by other exposures that are likely to occur in women exposed to CT, which are also capable of causing reproductive damage (38). Previously, we attempted to derive estimates from a Dutch case-control study (39), taking account of the sensitivity and specificity of assays. That study was based on a form of the "2-3" mixture model, but utilized reported summary data which did not allow titer distributions to be modeled. However, according to its authors, recruitment to the original study was likely to be subject to selection biases (40), and the resulting estimate of 45% (95% CrI: 28, 62) is likely to be an over-estimate. An estimate of 64% in Scotland was described as an upper bound (41). All estimates of the PEF are, of course, specific to time and place. For public health purposes, estimates should be based on contemporary local data.

This study shows how serum antibody titer distributions from case-control studies can be used to generate estimates of the PEF, based on finite mixture analysis. By attributing the causal mechanisms for TFI to differences between cases and controls in specific components of the titer distributions, rather than to differences in the overall prevalence of antibody, the mixture modeling approach reduces the extent to which PEF estimates are vulnerable to confounding, although it does not eliminate it, as discussed below.

The demonstration that there are four component distributions might appear surprising, rather than the two-component +ves and –ves model that might have been expected. However, the source data used in this exercise (29) has features that were apparent in earlier literature. Histograms suggesting two CT+ve "peaks" in control series have been published previously (11, 12, 16). Similarly, the very high titers seen in women with reproductive damage have also been

well-documented for PID / salpingitis (10, 18), and TFI (11, 15, 16, 30, 31). Our analyses suggest the fourth *CT+++* component occurs only in TFI cases. We attempted to fit "4-4" models but were unable to achieve stable results. Possibly, evidence for a "4-4" model might be obtained with a larger sample, although our efforts to fit these models suggests that very few controls would be in the *CT+++* group.

Interpretation of the different antibody positive groups must be somewhat speculative. Given that salpingitis is a necessary condition for TFI (6, 32, 42), it seems reasonable to regard the *CT+++* group, which is observed in cases only, as representing a causal mechanism linked to CT-related TFI. It is tempting to attribute this to a greater inflammatory response possibly due to a higher infectious load in those who develop TFI, as has been observed at the lower genital tract and in PID (43, 44). The 28.0% (6.9, 50.0) estimate of PEF based on the *CT+++* distribution alone can be regarded as a lower bound because it ignores the excess *CT++* observed in the cases.

This excess was substantial: 29.4% of TFI cases were in the *CT++* group, compared to the 6.5% of the controls (Table 4). We may speculate that the *CT++* peak in the women without TFI might represent women who have had upper genital tract infection in whom inflammation has resolved, either following treatment or spontaneously, without causing tubal damage (6, 45), as well as women with recent lower genital tract infections or re-infections, as the decline in antibody over time is far less marked in second infections (46). The excess *CT++* seen in cases could simply be due to increased exposure to CT in women whose TFI was in fact caused by other sexually transmitted infections or bacterial vaginosis. Bacterial vaginosis is associated with TFI (47) and with increased likelihood of CT infection and PID (48, 49). Sexual activity may

also lead to ascending infection with common respiratory or enteric pathogens that colonize the genital tract, which are also capable of causing reproductive damage (50).

Alternatively, women with TFI are more likely to have been exposed to repeat CT infections, which is associated with both reproductive damage (44, 51) and with higher titers (46). For this reason we may regard the higher PEF estimate (43.0%, 95%CrI: 27.6, 57.5) as an upper bound as it ascribes the *entire* excess *CT++* in cases to a causal mechanism rather than being partly or wholly the result of positive confounding.

The advantage of the mixture model estimates compared to the standard formula for PEF from case-control studies (52) $PEF = [\pi_{CT}(OR-1)]/[\pi_{CT}(OR-1)+1]$ is that they use the titer distribution as a marker of causal effect. This does not remove vulnerability to confounding, but it does limit it to a proportion of the *CT++* distribution, subject of course to our interpretation of the distributions. We can contrast our estimates with those obtained from the standard formula for PEF from case control studies. Using the same WIF data with titres at 1:64 and below as negatives, the Odds Ratio for TFI from Table 1 is (380 x 358 / 76 x 193) = 9.27 (95% CI: 6.9, 12.6). A population-based survey of the prevalence of Chlamydia antibody in 16-24 year old women in England, 2007-2010 (53) generated an estimate of 22.9% (95% CrI: 20, 26) in 23-24 year olds. This is most likely an underestimate of π_{CT} in the case-control study because the mean age of women in the WIF data was 30. Applying the formula to these estimates gives a PEF of 65.4%, or 71.3% if π_{CT} is 30%. Both these estimates are upper bounds as they attribute *all* the excess prevalence to a causal effect; the lower bound is zero, representing the case

where CT has no causal role in TFI but exposure to CT is common in those exposed to the true causes.

Nevertheless, both higher and lower estimates generated by the mixture models should be viewed cautiously for two reasons. First, the modeling process was not all pre-planned: each successive model was data-driven, motivated by poor fit in the previous model. The formulae for the PEF estimates were also developed *post hoc*. Second, our findings are based on an analysis of data collected for a different purpose. Our results therefore need to be confirmed by a specifically planned study, using modern assays, many of which are far more specific (33). The method could also be extended by testing samples for evidence of other pathogens capable of causing reproductive damage, including *Mycoplasma genitalium (54),* Bacterial vaginosis (47) and possibly *Neisseria gonorrhoeae* (50). The exercise could also be carried out on samples from women with PID and EP. If our results can be confirmed, finite mixture modeling may offer a way of quantifying the role of *Chlamydia* in reproductive damage, and form the basis for monitoring the impact of CT control programs in the population over time.

<3712 words>

Table 1. Numbers of Samples From (a) TFI Cases and Controls Seen at Bristol Between 1985-1995 and (b) From Secondary Anonymized Samples submitted at Bristol Public Health Laboratories during 2013 According to WIF Titer.

Titer Group Category	WIF titer	(a) Case-control study Controls	Cases	(b) Secondary data Number Negative	Percent negative	Total
1	<1:64	380	76	150	89.3	168
2	1:64	61	26	10	55.6	18
3	1:128	45	34	2	20.0	10
4	1:256	28	33	4	8.2	49
5	1:512	20	48	0	0	39
6	1:1024	30	122	0	0	17
7	1:2048	9	69	-	-	-
8	1:4096	0	22	-	-	-
9	>1:4096	0	4	-	-	-
Totals		573	434	166		301

Abbreviations: WIF, whole cell inclusion immune-fluorescence, Pgp-3, an immunogenic protein secreted by Chlamydia Trachomatis.

Table 2. Component Distributions in the Alternative Models and Estimators of Population Excess Fraction								
	Controls			Cases				Estimator for Population Excess Fraction
Model	*CT-*	*CT+*	*CT++*	*CT-*	*CT+*	*CT++*[a]	*CT+++*[a]	
"2-3"	Y	Y		Y	Y	Y		$\pi_{1,CT++}$
"3-3"	Y	Y	Y	Y	Y	Y		$\pi_{1,CT++} - (\pi_{1,CT+} + \pi_{1,CT++})\left(\frac{\pi_{0,CT++}}{\pi_{0,CT+} + \pi_{0,CT++}}\right)$
"3-4"	Y	Y	Y	Y	Y	Y	Y	$\left[\pi_{1,CT+++},\ \pi_{1,CT+++} + \pi_{1,CT++} - (\pi_{1,CT+} + \pi_{1,CT++})\left(\frac{\pi_{0,CT++}}{\pi_{0,CT+} + \pi_{0,CT++}}\right)\right]$ [b]

Abbreviations: Model "2-3", the control samples are a mixture of two distributions and cases are a mixture of three distributions; Model "3-3", the control samples are a mixture of three distributions and cases are a mixture of three distributions; Model "3-4", the control samples are a mixture of three distributions and cases are a mixture of four distributions; CT-, Not infected Chlamydia Trachomatis ; CT+, Chlamydia Trachomatis previously infected but with no immune response; CT++, Chlamydia Trachomatis previously infected with immune response; CT+++, Chlamydia Trachomatis previously infected with exceptionally high levels of serum antibody.

[a] Latent group labels should be thought as mnemonics (see Models in Methods section),

[b] [,]: [Minimum, Maximum]

Table 3. Observed and Predicted Frequency Counts of Each Titer, for Each Model, and Residual Deviance.

Group	Titers	Observed	Model "2-3"		Model "3-3"		Model "3-4"		"3-4" with Secondary data[a]	
			Predicted	Residual Deviance	Predicted	Residual Deviance	Predicted	Residual Deviance	Predicted	Residual Deviance
Controls[b]	<1:64	380	378.1	1.0	378.0	1.0	378.3	1.0	377.8	1.0
	1:64	61	62.2	1.0	60.9	1.0	60.5	0.9	63.3	0.8
	1:128	45	42.2	0.7	44.9	0.5	45.4	0.5	42.9	0.5
	1:256	28	34.1	1.6	29.7	0.7	29.3	0.6	28.4	0.5
	1:512	20	26.6[c]	2.1[c]	21.4	0.5	19.6	0.6	20.7	0.5
	1:1024	30	17.0[c]	8.9[c]	21.3[c]	4.0[c]	28.8	1.0	28.8	1.0
	1:2048	9	8.4	0.4	13.1	1.8	10.3	0.8	10.1	0.8
	1:4096	0	3.2	6.4	3.3	6.6	0.7	1.4	0.8	1.6
	>1:4096	0	1.2	2.4	0.4	0.9	0.1	0.1	0.1	0.2
	Total	573		25.4		17.0		7.1		6.8
TFI[b]	<1:64	76	77.4	1.0	76.6	1.0	76.4	1.0	76.1	1.0
	1:64	26	25.6	0.5	29.5	1.0	29.0	1.0	28.5	0.6
	1:128	34	29.8	1.1	30.5	1.0	31.8	0.7	32.2	0.6
	1:256	33	35.5	0.6	31.0	0.8	31.2	0.9	31.8	0.7
	1:512	48	57.9[c]	2.6[c]	59.7[c]	3.4[c]	52.7	1.4	51.8	1.2
	1:1024	122	105.1[c]	4.2[c]	108.1[c]	2.9[c]	118.4	1.1	118.5	1.1
	1:2048	69	80.0[c]	2.5[c]	77.1	1.6	69.3	0.7	69.7	0.8
	1:4096	22	20.0	1.2	19.3	1.1	20.1	1.0	20.5	0.9
	>1:4096	4	2.7	1.1	2.2	2.2	5.2	1.1	5.0	1.0
	Total	434		14.7		15.1		8.8		7.8
	Grand Total			39.2		32.1		15.9		14.6

Abbreviations: TFI, tubal factor infertility; Model "2-3", the control samples are a mixture of two distributions and cases are a mixture of three distributions; Model "3-3",the control samples are a mixture of three distributions and cases are a mixture of three distributions; Model "3-4", the control samples are a mixture of three distributions and cases are a mixture of four distribution

[a] Anonymised samples submitted for infertility investigations at the Bristol Public Health Laboratory during 2013

[b] Women undergoing infertility investigations at the Reproductive Medicine Clinic, at St. Michael's hospital, Bristol, during 1985-1995.

[c] Poorly fitting observations.

Table 4. Posterior Summaries From the 4 Models: Mean and Standard Deviation of the log Titers, and Percent in Each Component d in the Controls, π_{0d}, and TFI Cases π_{1d} along with Estimates of the Population Excess Fraction.

	Model "2-3"		Model "3-3"		Model "3-4"		Model "3-4" with secondary data[a]	
Group Means μ_d	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI	Mean	95% CrI
$\boldsymbol{\mu_1}$	-0.61	-1.6, -0.00	-1.09	-3.5,-0.01	-1.15	-3.5, -0.03	-0.70	-1.4, -0.2
$\boldsymbol{\mu_2}$	3.48	2.3, 4.5	2.41	1.3, 4.5	2.28	1.1, 3.4	2.72	2.1, 3.4
$\boldsymbol{\mu_3}$	5.84	5.7, 6.0	5.72	5.5, 5.9	5.60	5.3, 5.8	5.59	5.3, 5.8
$\boldsymbol{\mu_4}$					5.90	4.9, 7.0	6.10	5.2, 7.2
Group Standard deviations σ_d								
$\boldsymbol{\sigma_1}$	1.55	0.8, 2.2	1.54	0.51, 2.3	1.48	0.48, 2.2	1.60	1.2, 2.0
$\boldsymbol{\sigma_2}$	1.87	1.5, 2.3	1.66	1.0, 2.2	1.55	1.0, 2.1	1.54	1.2, 1.9
$\boldsymbol{\sigma_3}$	0.83	0.6, 1.0	0.90	0.60, 1.1	0.56	0.3, 0.8	0.56	0.3,0.8
$\boldsymbol{\sigma_4}$					1.19	0.7, 1.9	1.11	0.7, 1.8
Mixing proportions: Controls[b]								
$\boldsymbol{\pi_{01}}$	74.5	57, 86	67.9	45, 91	66.5	41, 87	74.0	66, 83
$\boldsymbol{\pi_{02}}$	25.5	14, 43	25.4	1.8, 48	26.6	5.8, 52	19.4	10, 28
$\boldsymbol{\pi_{03}}$			6.65	2.6, 11	6.92	3.7, 10	6.54	3.3, 10
Mixing proportions: Cases[b]								
$\boldsymbol{\pi_{11}}$	16.8	6.3, 26	12.7	1.3, 25	13.1	1.4, 24	16.4	9.7, 22
$\boldsymbol{\pi_{12}}$	35.5	25, 47	30.0	17, 44	26.9	15, 40	26.3	15, 37
$\boldsymbol{\pi_{13}}$	47.8	34, 58	57.3	35, 66	28.3	8.7,50	29.4	9.5, 50
$\boldsymbol{\pi_{14}}$	-	-		-	31.7	8.7, 53	28.0	6.9, 50
Population Excess Fraction	47.7	34.2, 57.8	35.9	-21.3, 52.7	31.7[c]	8.7, 52.8	28.0[a]	6.9, 50.0
					46.8[d]	23.2, 64.1	43.0[b]	27.6, 57.5

Abbreviations: TFI, Tubal factor infertility; CrI, Credible interval; Model "2-3", the control samples are a mixture of two distributions and cases are a mixture of three distributions; Model "3-3", the control samples are a mixture of three distributions and cases are a mixture of three distributions; Model "3-4", the control samples are a mixture of three distributions and cases are a mixture of four distributions.

[a] Anonymised samples submitted for infertility investigations at the Bristol Public Health Laboratory during 2013

[b] Women undergoing infertility investigations at the Reproductive Medicine Clinic, at St. Michael's hospital, Bristol, during 1985-1995.

[c] Lower bound/minimum estimate for the population excess fraction

[d] Upper bound/maximum estimate for the population excess fraction

Acknowledgements

Author affiliations: School of Social and Community Medicine, University of Bristol, Bristol, UK (AE. Ades, D. Kounali, PJ. Homer); Bristol Sexual Health Centre, University Hospitals Bristol NHS Trust, Bristol, UK (PJ. Homer); Institute of Applied Health Research, University of Birmingham, Birmingham, UK (MJ. Price); Jefferiss Trust Laboratories, Wright-Fleming Institute, Imperial College London, London, UK (GS. Wills, MO., McClure); Public Health Laboratory Bristol, National Infection Service, Public Health England, Myrtle Road, Bristol, UK. (P. Muir); School of Clinical Sciences, University of Bristol and Bristol Centre for Reproductive Medicine, North Bristol NHS Trust, Bristol, UK (VA. Akande).

This work was supported by the NIHR Health Protection Research Unit in Evaluation of Interventions at the University of Bristol under award number HPRU-2012-10026 and in partnership with Public Health England (PHE). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

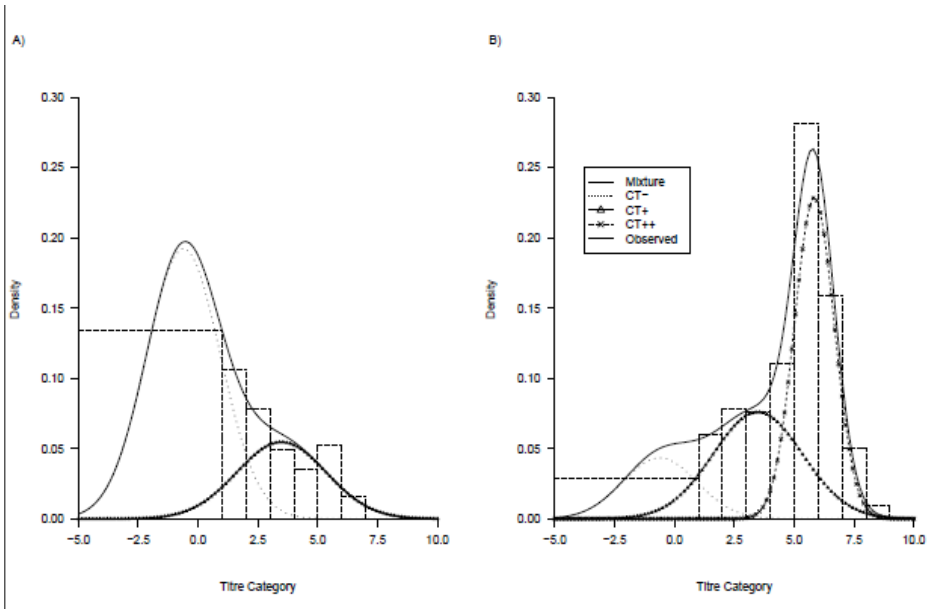
Conflict of interest: none declared.

References

1. Price M, Ades A, De Angelis D, et al. Risk of Pelvic Inflammatory Disease following *Chlamydia trachomatis* infection: analysis of prospective studies with a multistate model. *Am J Epidemiol* 2013;178(3):484-492.
2. Gottlieb SL, Martin DH, Xu FJ, et al. Summary: the natural history and immunobiology of Chlamydia trachomatis genital infection and implications for Chlamydia control. *Journal of Infectious Diseases* 2010;201:S190-S204.
3. Haggerty C, Gottlieb S, Taylor B, et al. Risk of sequelae after Chlamydia trachomatis genital infection in women. *The Journal of Infectious Diseases* 2010;201(S2):134-155.
4. Wallace LA, Scouler A, Hart G. What is the excess risk of infertility in women after genital chlamydia infection? A systematic review of the evidence. *Sexually Transmitted Infections* 2008;84(3):171-175.
5. Westrom L. Effect of pelvic inflammatory disease on fertility. *Venereology-the Interdisciplinary International Journal of Sexual Health* 1995;8(4):219-222.
6. Westrom L, Joesoef R, Reynolds G, et al. Pelvic inflammatory disease and fertility - a cohort study of 1,844 women with laparoscopically verified disease and 657 control women with normal laparoscopic results. *Sexually Transmitted Diseases* 1992;19(4):185-192.
7. Westrom L, Bengtsson LP, Mardh PA. Incidence, trends, and risks of ectopic pregnancy in a population of women. *British Medical Journal* 1981;282(6257):15-18.
8. Wolner-Hanssen P. Silent pelvic inflammatory disease - is it overstated? *Obstetrics and Gynecology* 1995;86(3):321-325.
9. Ripa KT, Svensson L, Treharne JD, et al. Chlamydia-Trachomatis Infection in Patients with Laparoscopically Verified Acute Salpingitis - Results of Isolation and Antibody Determinations. *American Journal of Obstetrics and Gynecology* 1980;138(7):960-964.
10. Treharne JD, Ripa KT, Mardh PA, et al. Antibodies to Chlamydia-Trachomatis in Acute Salpingitis. *British Journal of Venereal Diseases* 1979;55(1):26-29.
11. Punnonen R, Terho P, Nikkanen V, et al. Chlamydial Serology in Infertile Women by Immunofluorescence. *Fertility and Sterility* 1979;31(6):656-659.
12. Conway D, Glazener CM, Caul EO, et al. Chlamydial serology in fertile and infertile women. *Lancet* 1984;1(8370):191-193.
13. Kosseim M, Brunham RC. Fallopian tube obstruction as a sequela to *Chlamydia trachomatis* infection. *European Journal of Clinical Microbiology* 1986;5(5):584-590.
14. Persson K, Osser S, Birkelund S, et al. Antibodies to Chlamydia trachomatis heat shock proteins in women with tubal factor infertility are associated with prior infection by C-trachomatis but not by C-pneumoniae. *Human Reproduction* 1999;14(8):1969-1973.
15. Peek JC, Graham FM, Hookham A. Prevalence of Chlamydial Antibodies in Women with Tubal Disease - Impact of Chlamydia Trachomatis on the Demand for Invitro Fertilization. *New Zealand Medical Journal* 1990;103(884):63-65.
16. Anestad G, Lunde O, Moen M, et al. Infertility and Chlamydial Infection. *Fertility and Sterility* 1987;48(5):787-790.
17. Kihlstrom E, Lindgren R, Ryden G. Antibodies to *Chlamydia trachomatis* in women with infertility, pelvic inflammatory disease and ectopic pregnancy. *European Journal of Obstetrics Gynecology and Reproductive Biology* 1990;35(2-3):199-204.
18. Gump DW, Gibson M, Ashikaga T. Evidence of Prior Pelvic Inflammatory Disease and Its Relationship to Chlamydia-Trachomatis Antibody and Intrauterine Contraceptive Device Use in Infertile Women. *American Journal of Obstetrics and Gynecology* 1983;146(2):153-159.
19. Sellors JW, Mahony JB, Chernesky MA, et al. Tubal factor infertility - an association with prior chlamydial Infection and asymptomatic salpingitis. *Fertility and Sterility* 1988;49(3):451-457.

20. Paavonen J, Westrom L, Eschenbach DA. Pelvic Inflammatroy Disease. In: Holmes K, Sparling PF, Stamm WE, eds. *Sexually Transmitted Disease*. London: McGraw Hill, 2008:1021-1022.
21. Persson K. The role of serology, antibiotic susceptibility testing and serovar determination in genital chlamydial infections. *Best Practice & Research in Clinical Obstetrics & Gynecology* 2002;16(6):801-814.
22. Johnson AM, Horner P. A new role for Chlamydia trachomatis serology? *Sexually Transmitted Infections* 2008;84(2):79-80.
23. McLachlan GJ, Peel D. *Finite Mixture Models*. New York: Wiley; 2000.
24. Baughman AL, Bisgard KM, Lynn F, et al. Mixture model analysis for establishing a diagnostic cut-off point for pertussis antibody levels. *Stat Med* 2005;25(17):2994-3010.
25. Pai M, Dendukuri N, Wang L, et al. Improving the estimation of tuberculosis infection prevalence using T-cell-based assay and mixture models. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2008;12(8):895-902.
26. Trebucq A, Guerin N, Ali Ismael H, et al. Prevalence and trends of infection with Mycobacterium tuberculosis in Djibouti, testing an alternative method. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2005;9(10):1097-1104.
27. Neuenschwander BE, Zwahlen M, Kim SJ, et al. Determination of the prevalence of infection with Mycobacterium tuberculosis among persons vaccinated against Bacillus Calmette-Guerin in South Korea. *Am J Epidemiol* 2002;155(7):654-663.
28. Hadgu A, Dendukuri N, Hilden J. Evaluation of nucleic acid amplification tests in the absence of a perfect gold-standard test: a review of the statistical and epidemiologic issues. *Epidemiology (Cambridge, Mass)* 2005;16(5):604-612.
29. Akande VA, Hunt LP, Cahill DJ, et al. Tubal damage in infertile women: prediction using chlamydia serology. *Human Reproduction* 2003;18(9):1841-1847.
30. Meikle SF, Zhang XZ, Marine WM, et al. Chlamydia-Trachomatis Antibody-Titers and Hysterosalpingography in Predicting Tubal Disease in Infertility Patients. *Fertility and Sterility* 1994;62(2):305-312.
31. Robertson JN, Ward ME, Conway D, et al. Chlamydial and Gonococcal Antibodies in Sera of Infertile Women with Tubal Obstruction. *Journal Of Clinical Pathology* 1987;40(4):377-383.
32. Westrom L, Wolner-Hanssen P. Pathogenesis of pelvic inflammatory disease. *Genitourinary Medicine* 1993;69(1):9-17.
33. Wills GS, Horner PJ, Reynolds R, et al. Pgp3 antibody Enzyme-Linked Immunosorbent Assay, a sensitive and specific assay for seroepidemiological analysis of Chlamydia trachomatis Infection. *Clinical and Vaccine Immunology* 2009;16(6):835-843.
34. Spiegelhalter DJ, Best NG, Carlin BP, et al. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society (B)* 2002;64(4):583-616.
35. McCullagh P, Nelder JA. *Generalized Linear Models, Second Edition*. Taylor & Francis; 1989.
36. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 1998;7(4):434-455.
37. Land JA, van Bergen JEAM, Morre SA, et al. Epidemiology of Chlamydia trachomatis infection in women and the cost-effectiveness of screening. *Human Reproduction Update* 2010;16(2):189-204.
38. Taylor-Robinson D, Jensen JS, Svenstrup HF, et al. Difficulties experienced in defining the microbial cause of pelvic inflammatory disease. *International Journal of STD AIDS* 2012;23(1):18-24.
39. Price MJ, Ades AE, Welton NJ, et al. How much Tubal Factor Infertiliy is caused by chlamydia? Estimation based on serological evidence adjusted for sensitivity and specificity. *Journal of Sexually Transmitted Diseases* 2011;39(8):608-613.
40. Land JA, Gijsen AP, Kessels AGH, et al. Performance of five serological chlamydia antibody tests in subfertile women. *Human Reproduction* 2003;18(12):2621-2627.

41. Kavanagh K, Wallace LA, Robertson C, et al. Estimation of the risk of tubal factor infertility associated with genital chlamydia infection in women: a statistical modelling study. *International Journal Of Epidemiology* 2013;42(2):493-503.
42. Cahill DJ, Wardle PG. Management of infertility. *British Medical Journal* 2002;325(7354):28-32.
43. Michel C-EC, Sonnex C, Carne CA, et al. Chlamydia trachomatis load at matched anatomic sites: Implications for screening strategies. *Journal of Clinical Microbiology* 2007;45(5):1395-1402.
44. Geisler WM, Suchland RJ, Whittington WLH, et al. Quantitative culture of Chlamydia trachomatis: Relationship of inclusion-forming units produced in culture to clinical manifestations and acute inflammation in urogenital disease. *Journal of Infectious Diseases* 2001;184(10):1350-1354.
45. Tait IA, Duthie SJ, Taylor-Robinson D. Silent upper genital tract chlamydial infection and disease in women. *International Journal of Std & Aids* 1997;8(5):329-331.
46. Horner PJ, Wills GS, Reynolds R, et al. Effect of time since exposure to Chlamydia trachomatis on chlamydia antibody detection in women: a cross-sectional study. *Sexually Transmitted Infections* 2013;89(5):398-403.
47. van Oostrum N, De Sutter P, Meys S, et al. Risks associated with bacterial vaginosis in infertility patients: a systematic review and meta-analysis. *Human Reproduction* 2013;28(7):2039-2048.
48. Taylor B, Darville T, Haggerty C. Does bacterial vaginosis cause pelvic inflammatory disease? *Sexually Transmitted Diseases* 2013;40(2):117-122.
49. Allsworth JE, Peipert JF. Severity of bacterial vaginosis and the risk of sexually transmitted infection. *American Journal of Obstetrics & Gynecology* 2011;205(2):113-116.
50. Brunham RC, Gottlieb SL, Paavonen J. Pelvic inflammatory disease. *New England Journal of Medicine* 2015;372(21):2039-2048.
51. Hillis SD, Owens LM, Marchbanks PA, et al. Recurrent chlamydial infections increase the risks of hospitalization for ectopic pregnancy and pelvic inflammatory disease. *American Journal of Obstetrics & Gynecology* 1997;176(1 Pt 1):103-107.
52. Rockhill B, Newmam B, Weinberg C. Use and misuse of population attributable fractions. *American Journal of Public Health* 1998;88(1):15-19.
53. Horner P, Soldan K, Vieira SM, et al. C. trachomatis pgp3 Antibody Prevalence in Young Women in England, 1993–2010. *Plos One* 2013;8(8).
54. Haggerty CL, Taylor BD. Mycoplasma genitalium: an emerging cause of pelvic inflammatory disease [electronic article, review]. *Infectious diseases in obstetrics and gynecology.* Document ID : 959816, Vol. 2011, 9 pages (DOI:2011. doi:10.1155/2011/959816).



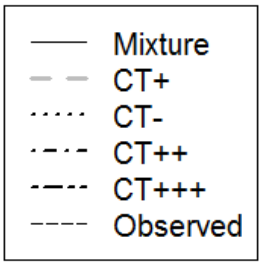
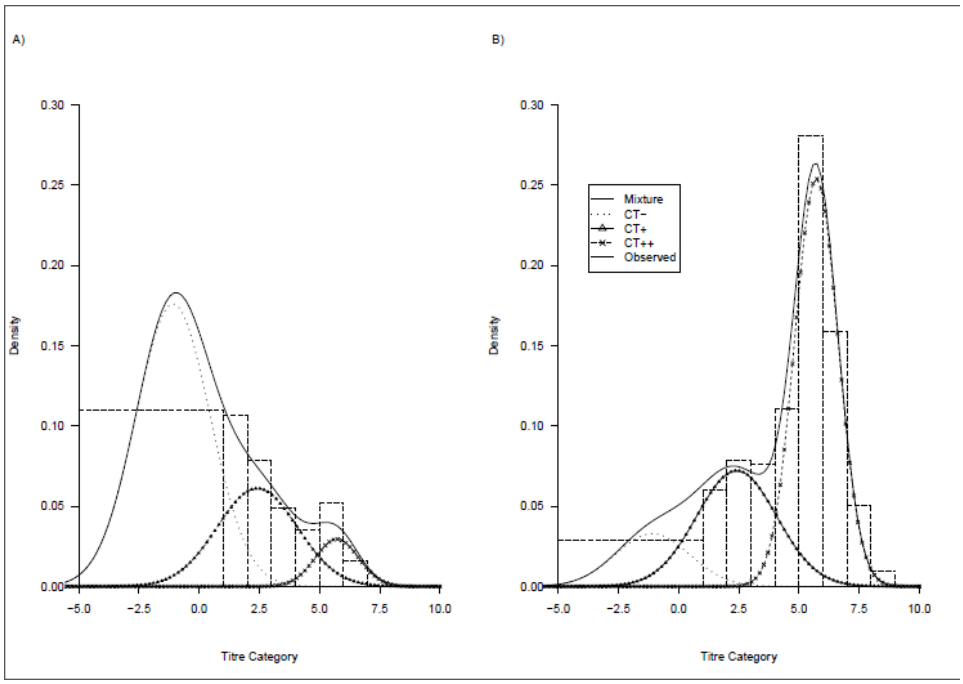


Figure 1. Fitted component distributions for controls (Panel A) and cases (Panel B) undergoing infertility investigations at the Reproductive Medicine Clinic, at St. Michael's hospital, Bristol, during 1985-1995, based on the model assuming two component distributions for cases and three for controls[a].

[a] The figure is drawn to a scale reflecting the mixing proportions, the predicted overall titer distribution (solid black line), and the observed data (dashed line histogram). The left-most



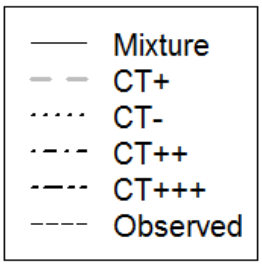


Figure 2. Fitted component distributions for controls (Panel A) and cases (Panel B) undergoing infertility investigations at the Reproductive Medicine Clinic, at St. Michael's hospital, Bristol, during 1985-1995, based on the model assuming three component distributions for cases and three for controls[a].

histogram bar comprises titers below 1:64. This has been plotted to cover the area -5 to +1 on the log titer scale; its area corresponds to the proportion of data at these titers.

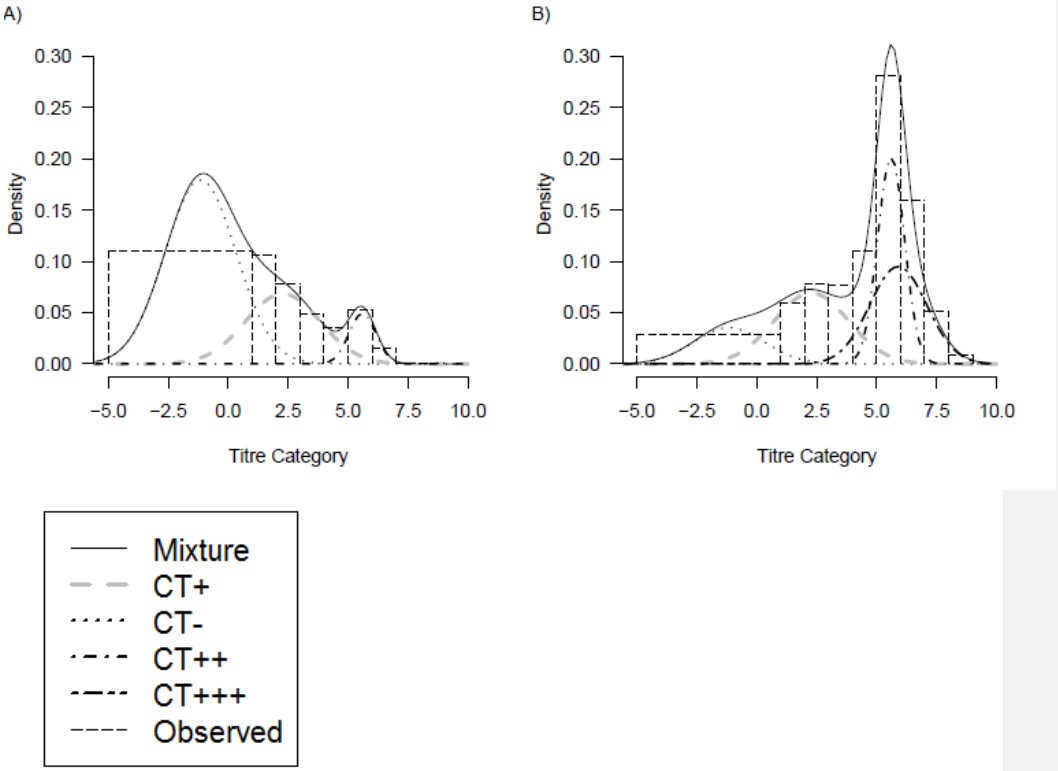


Figure 3. Fitted component distributions for controls (Panel A) and cases (Panel B) Figure 1. Fitted component distributions for Controls (Panel A) and cases (Panel B) undergoing infertility investigations at the Reproductive Medicine Clinic, at St. Michael's hospital, Bristol, during 1985-1995, based on the model assuming three component distributions for cases and four for controls[a].

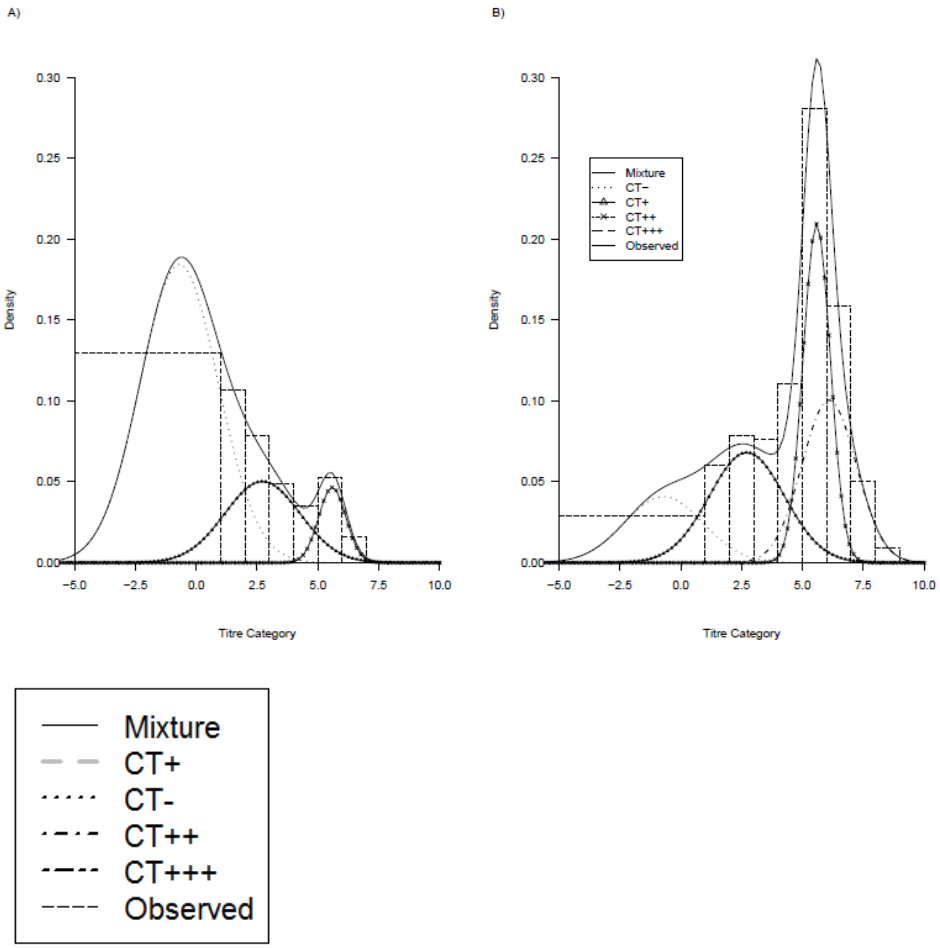


Figure 4. Fitted component distributions for Controls (Panel A) and cases (Panel B) undergoing infertility investigations at the Reproductive Medicine Clinic, at St. Michael's hospital, Bristol, during 1985-1995, based on the model assuming three component distributions for cases and four for controls and which also made use of secondary data on Anonymised samples submitted for infertility investigations at the Bristol Public Health Laboratory during 2013.